

CIRCULAR DICHROISM AND SEQUENCE OF NUCLEOTIDES IN NUCLEIC ACIDS

G.B. Zavil'gel'skiy, T.V. Venkstern, A.A. Bayev

N66-23543

FACILITY FORM 602

(ACCESSION NUMBER)

(THRU)

(PAGES)

(CODE)

(NASA CR OR TMX OR AD NUMBER)

(CATEGORY)

Translation of "Tsirkulyarnyy dikhroizm i posledovatel'nost'
nukleotidov v nukleinovyykh kislotakh".
Doklady Akademii Nauk SSSR, Vol. 166, No. 4, pp. 978-981, 1966

GPO PRICE \$ _____

CFSTI PRICE(S) \$ _____

Hard copy (HC) 1.00Microfiche (MF) .50

ff 653 July 65

NATIONAL AERONAUTICS AND SPACE ADMINISTRATION
WASHINGTON D.C. MARCH 1966

NASA TT F-10,049

CIRCULAR DICHROISM AND SEQUENCE OF NUCLEOTIDES IN NUCLEIC ACIDS

(Presented by Academician V. A. Engel'gardt, October 16, 1965)

G. B. Zavil'gel'skiy, T. V. Venkstern, A. A. Bayev

ABSTRACT

23543

This article proposes to use circular dichroism to determine the order of the bases in oligonucleotides, and finds that this method is effective.

The procedure for determining the sequence of bases in oligonucleo- /978* tides which is carried out, for example, when studying the primary structure of nucleic acids is laborious and not always reliable. This is equally true of its chemical and enzymatic variations. Spectrophotometry (Ref. 1) enables us to determine the composition of di- and trinucleotides, but is powerless to establish the sequence of bases in them. It is natural that the desire should arise to make analysis of oligonucleotides automatic and to develop new, particularly physical, methods of ascertaining sequence.

Measurement of circular dichroism (c.d.) in synthetic polynucleotides, RNA, and DNA has shown that in the 220-300 mμ region the shape and intensity of dichroic bands in polymers is the simple sum of the contributions of the individual nucleotides and may serve as a sensitive test of changes in secondary structure (Ref. 2-4). The assumption has been expressed that

* Note: Numbers in the margin indicate pagination in the original foreign text.

the c.d. effect and, consequently, anomalous dispersion of optical activity (a.d.o.a.) in polymers, is principally determined by the reaction of adjacent nitrogen bases. Theoretical computation of the form of c. d. curves for spiral polynucleotides agreed well with experiment (Ref. 5). The same work advanced the hypothesis that the c.d. effect determined by the reaction between bases should be observed even in the simplest system composed of a nucleotide pair. Measurement of the a.d.o.a. curve (and of the c.d. curve recalculated by the Kronig-Kramers conversion) in the dinucleotide ApA completely confirmed the validity of the theoretical assumptions (Ref. 6).

The present work proposes to use c.d. to determine the order of the bases in oligonucleotides. In so doing we start from two positions:

(1) the c.d. effect in the dinucleotides is determined by the reaction between the bases and does not coincide with the sum of the c.d. effects of the separate nucleotides, and (2) in more complex oligonucleotides the main contribution to the c.d. curve is made by the reactions between the closest neighbors, i.e., for an oligonucleotide, the c.d. curve may be regarded in the first approximation as the sum of the c.d. curves of the individual nucleotide pairs.

The oligonucleotides were separated from the enzymatic hydrolysates of transport RNA of baker's yeast by means of ion-exchange chromatography and partition chromatography on paper; their composition and order were determined by methods described in (Ref. 7). The dinucleotide TpT was kindly furnished by D. Shugar (Institute of Biochemistry and Biophysics of the Polish Academy of Sciences, Warsaw) and the polyuridylic acid by R. S. Nezlin (Institute of Molecular Biology of the Academy of Science, USSR).

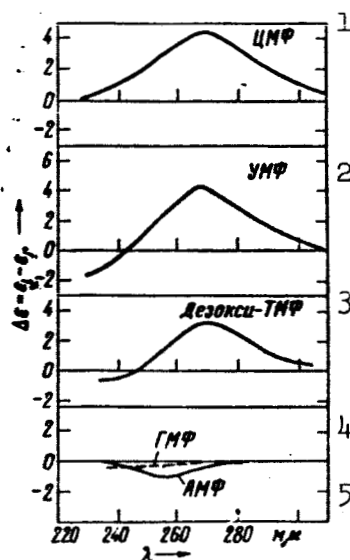


Figure 1

Circular Dichroism Curves of Mononucleotides. (All preparations were dissolved in distilled water [pH 5-6].)

1 - cytidine monophosphate; 2 - uridine monophosphate; 3 - deoxy-thymidine monophosphate; 4 - guanosine monophosphate; 5 - adenosine monophosphate.

Circular dichroism was measured on a dichrograph made by the French firm of Roussell-Jouan with three scales of sensitivity of 3, 2, and $1.5 \cdot 10^{-4}$ optical units per mm at room temperature. The mono- and oligonucleotides were dissolved in distilled water (pH 5-6); optical density of the solution was no more than 1.0. A pH of 1.0 was produced by adding 0.2 ml of 1 N HCl to 2.0 ml of the solution. The region in which circular dichroism was measured was 230-310 mμ.

The following nucleotides and oligonucleotides were employed: 5'-adenosine, 5'-guanosine monophosphorate, deoxy-5'-thymidine phosphate; ApCp, ApGp, GpCp, CpGp, CpG, GpUp, TpT; CpApGp, mixture (50% CpApGp - 50% ApCpGp), UpUpUp, UpUpUpGp, ApCpApCpGp, polyuridylic acid.

Results

Figure 1 gives the c.d. curves of various mononucleotides. The pyrimidines in the long-wave region show a rather wide positive band with a /979 maximum at about 268-273 m μ . In the short-wave region (below 240 m μ) uridine monophosphate and thymidine monophosphate have a weak negative band with its maximum apparently below 220 m μ . The c.d. curves of the purines are considerably less pronounced than those of the pyrimidines and have a single negative band - in adenosine monophosphate with λ_{\max} about 255 m; the position of a weak maximum was not determined for guanosine monophosphate.

The c.d. curves of dinucleotides are as a rule of complex shape with several bands and, as analysis shows, are not the result of a simple addition of the c.d. of individual nucleotides (Figure 2). Instead of a small negative maximum, which is expected in the summation of the c.d. of nucleotides, for ApGp at pH 5-6, three bands are seen: (1) positive with $\lambda_{\max}^1 = 278$ m μ , (2) negative with $\lambda_{\max}^2 = 260$ m μ , and (3) positive with $\lambda_{\max}^3 < 220$ m μ . The c.d. curve of ApCp is asymmetrical with a very strong positive band having $\lambda_{\max} = 275$ m μ and a weak negative band having λ_{\max} around 235 m μ . Two clear bands of differing sign are seen in TpT ($\lambda_{\max}^1 = 277$ m μ and $\lambda_{\max}^2 = 250$ m μ). There are also two bands in GpUp, but both are positive - a very weak band with $\lambda_{\max} = 280$ m μ and a strong one with $\lambda_{\max} = 250$ m μ . The example of the G and C pair is interesting. In dinucleotide CpGp (and also CpG) the c.d. curve has a characteristic positive band with $\lambda_{\max} = 280$ m μ . The dinucleotide with the reverse order of bases, GpCp, has a broad positive band, which both in size and in location of the maximum (270 m μ) is congruent with the c.d. band for cytidine monophosphate (Figure 1). Therefore, in GpCp, the

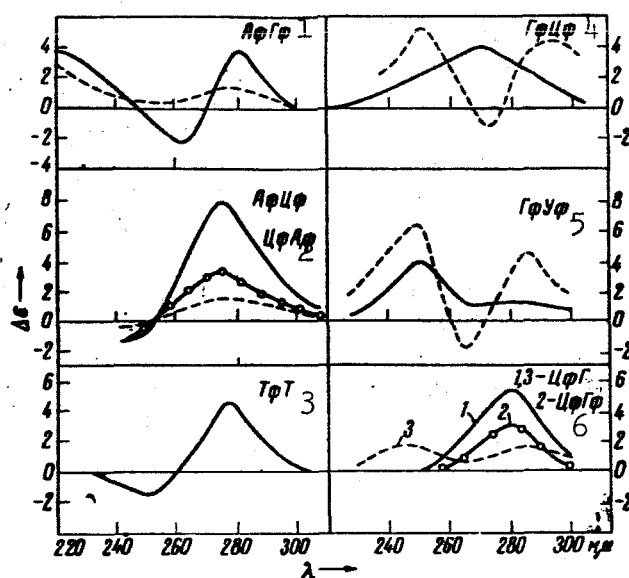


Figure 2

Circular Dichroism Curves

Solid curves - distilled water (pH 5-6). Broken curves - pH 1.0.

Circular dichroism curve for CpAp calculated from the trinucleotide CpApGp curve (Figure 3).

1 - ApGp; 2 - ApCp, CpAp; 3 - TpT; 4 - GpCp; 5 - GpUp; 6 - 1,3-CpG, 2-CpGp.

the effect of the individual bases is additive (with a pH of 5-6).

With a pH of 1.0 the c.d. curve of dinucleotides has its own shape. In ApGp, ApCp, and CpG the c.d. effect is sharply decreased, while in CpG there additionally appears a very weak positive band at 240-250 mμ. In GpCp and GpUp, on the contrary, the c.d. effect increases with a pH of 1.0. The /980 c.d. curves for these dinucleotides take on a characteristic form with three bands, two positive and one negative, which differ from each other only in the positions of the first two maxima - in GpCp $\lambda_{\max}^1 = 295 \text{ m}\mu$, $\lambda_{\max}^2 = 273 \text{ m}\mu$, and $\lambda_{\max}^3 = 250 \text{ m}\mu$, while in GpUp they are, respectively, 285 mμ, 265 mμ, and 250 mμ.

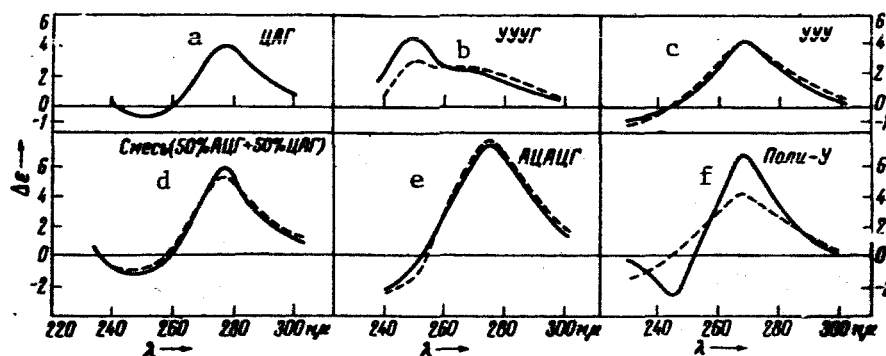


Figure 3

Circular Dichroism Curves of Oligonucleotides

Solid curves - experimental values; broken curves - theoretical values obtained in "additive pair" approximation. In the case of polyuridylic acid and UpUpUp the broken curve represents the c.d. of uridine monophosphate. All preparations were dissolved in distilled water (pH 5-6).

a - CAG; b - UUUG, c - UUU, d - Mixture (50% ACG - 50% CAG); e - AGACG; f - Polyuridylic acid.

The findings on the mono- and dinucleotides permit the following conclusions: (1) the c.d. effect of dinucleotides is as a rule not the result of adding the c.d. effects of the individual mononucleotides, and thus cannot serve as a measure of the specific reaction between the nucleotides; (2) the shape of the c.d. curve is characteristic for each pair, which makes it possible to determine easily from the c.d. curve the composition and sequence of the nucleotide in a pair (e.g., GpCp or CpGp); (3) the change in pH is a convenient additional factor for identifying the pair: pairs like ApXp (X - any nucleotide) have high c.d. in neutral and subacid pH, while pairs of the GpXp type have high c.d. with strongly acid pH.

If we assume that the characteristic c.d. curve shape occurs because of the specific reaction of the most closely adjoining nucleotides, there then arises the alluring possibility of using the c.d. method to determine

sequence in small oligonucleotides. In this case the c.d. curve of the oligonucleotide may be represented as the sum of the c.d. curves of the separate pairs encountered in the oligonucleotide in question (we will call this approximation "pair additivity"). The trinucleotide ApCpGp, for example, consists of the two pairs ApCp and CpGp, but we disregard the reaction between A and G.

To check the correctness of "additive pair" approximation, we measured the c.d. curves of several oligonucleotides (Figure 3).

Two bands are seen in the c.d. curve of CpApGp - a strong positive band with $\lambda_{\max} = 277 \text{ m}\mu$ and a weak negative band with $\lambda_{\max} = 245 \text{ m}\mu$. The zero line is crossed at $\lambda = 260 \text{ m}\mu$. In the "additive pair" approximation, the c.d. effect for this trinucleotide must be

$$\Delta\epsilon_{\text{CAG}} = \frac{2\Delta\epsilon_{\text{CA}} + 2\Delta\epsilon_{\text{AG}} - \Delta\epsilon_{\text{AMP}}}{3} \quad (1)$$

($\Delta\epsilon$ is the difference in molar extinction ($\epsilon_1 - \epsilon_r$) for one nucleotide).

Not having the pair CpAp, we used the c.d. curve for CpApGp and Formula (1) to compute $\Delta\epsilon_{\text{CA}}$. In Figure 2 the thus computed c.d. curve /981 for CpAp is depicted on the graph, together with the experimentally derived c.d. curve for ApCp - the curves were the same in shape, but the magnitude of the effect for CpAp is somewhat lower.

The c.d. curve for the mixture (50% CpApGp - 50% ApCpGp), as well as for CpApGp, displays two bands - a strong positive band with $\lambda_{\max} = 277 \text{ m}\mu$ and a weak negative one with $\lambda_{\max} = 245 \text{ m}\mu$. The magnitude of the effect, however, is somewhat higher. A similar situation is seen also for the pentanucleotide ApCpApCpGp, but the negative band maximum is slightly shifted toward the short waves ($\lambda < 240 \text{ m}\mu$). In the "additive pair" approximation,

the $\Delta\epsilon$ value for the mixture (ApCpGp - CpApGp) is determined from the formula

$$\Delta\epsilon_{\text{mixture}} = \frac{\Delta\epsilon_{AC} + \Delta\epsilon_{CA} + \Delta\epsilon_{CG} + \Delta\epsilon_{AG} - 1/2\Delta\epsilon_{CMP} - 1/2\Delta\epsilon_{AMP}}{3}, \quad (2)$$

and for the pentane nucleotide ApCpApCpGp, by the formula

$$\Delta\epsilon_{ACACG} = \frac{4\Delta\epsilon_{AC} + 2\Delta\epsilon_{CA} + 2\Delta\epsilon_{CG} - 2\Delta\epsilon_{CMP} - \Delta\epsilon_{AMP}}{5}, \quad (3)$$

The computational results are given in Figure 3. A rather exact agreement with the experimental curves is detected.

With the trinucleotide UpUpUp the c.d. curve coincides with that for the mononucleotide uridine monophosphate, which indicates lack of reaction between adjacent uracil radicals in small oligonucleotides.

In the tetranucleotide UpUpUpGp, the value of $\Delta\epsilon$ has the form

$$\Delta\epsilon_{UUUG} = \frac{4\Delta\epsilon_{UU} + 2\Delta\epsilon_{UG} - 2\Delta\epsilon_{UMP}}{4}, \quad (4)$$

or, if we consider that $\Delta\epsilon_{UU} = \Delta\epsilon_{UMP}$ (from the data of trinucleotide UpUpUp), then

$$\Delta\epsilon_{UUUG} = \frac{2\Delta\epsilon_{UMP} + 2\Delta\epsilon_{UG}}{4}, \quad (5)$$

Having no UpGp, we computed (5) on the assumption that

$$\Delta\epsilon_{UG} \cong \Delta\epsilon_{GU} \quad (6)$$

Figure 3 gives the results of the computations. The shape of the theoretical curve is close to that of the experimental one - there are two positive bands with λ_{max}^1 around 270 m μ and $\lambda_{\text{max}}^2 = 250$ m μ . The curves differ slightly in magnitude of the effect. This is apparently the result of the additional approximation of (6) which we made during the calculations.

The c.d. curve for polyuridylic acid (in contrast to UpUpUp) differs perceptibly from that for UMP - the c.d. magnitude is higher, and moreover,

there is a clearly marked negative maximum at the short wavelengths ($\lambda_{\text{max}} = 245 \text{ m}\mu$). In polyuridylic acid the secondary structure of the polymer apparently adds its own influence to the c.d. effect (see also [Ref. 2]).

The above statements permit us to assume that there is an "additive pair" principle and to give an optimistic evaluation of the use of the c.d. method for determining the sequence of bases in small oligonucleotides.

The authors are grateful to V. A. Engel'hardt for valuable discussion of the work.

Note in proof. While preparing the article for printing, we have become acquainted with the article by Charles Cantor and J. Tinoco which proposes the use of an a.d.o.a. method for determining sequence in oligonucleotides (Ref. 8).

Institute of Molecular Biology
USSR Academy of Sciences

Received
10/7/65

REFERENCES

1. Bajew, A.A., Winkstern, T.W., Mirsabekow, A.D. Abhandlungen der Deutsch. Akad. Wissensch., Berlin, No. 6, 423, 1964.
2. Brahms, J. J. Am. Chem. Soc., 85, 3298, 1963.
3. Brahms, J., Mommaerts, W., Sadron, Ch. C. R. Acad. Sci., Paris, 258, 2203, 1964.
4. Brahms, J., Mommaerts, W. J. Molec. Biol., 10, 73, 1964.
5. Tinoco, J. J. Am. Chem. Soc., 86, 297, 1964.
6. Warshaw, M., Bush, C.A., Tinoco, J. Biochem. Biophys. Res. Commun., 18, 663, 1964.

7. Mirzabekov, A., Krutilina, A., Gorshkova, V.I., Bayev, A.A.
Biokhimiya, 29, 1158, 1964.
8. Cantor, Ch.R., Tinoco, J. J. Molec. Biol., 13, 65, 1965.

Scientific Translation Service
4849 Tocaloma Lane
La Canada, California